AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Genome Sequences of Eight Shiga Toxin-Producing *Escherichia coli* Strains Isolated from a Produce-Growing Region in California

Craig T. Parker,[a] Kerry K. Cooper,[b] Steven Huynh,[a] Timothy P. Smith,[c] James L. Bono,[c] Michael Cooley[a]

[a]Western Regional Research Center, Produce Safety and Microbiology Research Unit, Agricultural Research Service, U.S. Department of Agriculture, Albany, California, USA

[b]School of Animal and Comparative Biomedical Sciences, University of Arizona, Tucson, Arizona, USA

[c]Meat Safety and Quality Research Unit, Agricultural Research Service, U.S. Department of Agriculture, Clay Center, Nebraska, USA

ABSTRACT Produce contaminated with Shiga toxin-producing *Escherichia coli* (STEC) is a continuing source of foodborne illness in the United States. This report documents the complete genome sequences of eight STEC strains isolated from livestock and water samples taken from a major agricultural region for leafy greens in California.

Cases of foodborne illness due to Shiga toxin-producing *Escherichia coli* (STEC) are a global problem, with approximately 2.8 million cases of acute illness annually (1). Although STEC O157:H7 strains are the cause of the majority of these illnesses, non-O157 STEC infections have emerged as a public health problem in the United States and internationally (2–4). The presence of STEC adjacent to a raw produce production environment must be considered a risk factor for contamination, with several transmission routes possible, including water, air, and animals (5–9). Indeed, raw produce has been the source of multiple outbreaks involving a variety of identified serotypes, including O157, O145, and O104 STEC (10–14). This announcement documents the complete genome sequences of eight non-O157 STEC strains previously isolated from livestock and water samples collected in a major agricultural region for leafy greens production on California's Central Coast (15).

The STEC cultures were grown overnight on LB agar plates at 37°C, and genomic DNA was extracted using sucrose-Tris with phenol-chloroform cleanup extractions as described previously (16). Sequencing was performed using a Pacific Biosciences (PacBio, Menlo Park, CA) RS II platform with 20-kb SMRTbell libraries as described previously (17). Briefly, SMRTbell libraries were prepared from 10 μg of bacterial genomic DNA with G-tube (Covaris, Woburn, MA) fragmentation using the described PacBio procedure (18) but with 1× AMPure cleanup and DNA repair after 10-kb size selection using BluePippin. Single-molecule real-time (SMRT) cells were run with the 0.1 nM on-plate concentration, P6/C4 sequencing chemistry, MagBead One Cell Per Well v1 collection protocol, and 360-min data collection mode. The PacBio reads were assembled using the hierarchical genome assembly process (HGAP) v3.0 in SMRT Analysis v2.3.0 (PacBio). The chromosomes and plasmids for each STEC strain were circularized from one contig with overlapping ends. All strains were sequenced using an Illumina MiSeq platform and the KAPA LTP library preparation kit (KAPA Biosystems, Wilmington, MA). The pooled libraries were loaded into a MiSeq system and sequenced using a MiSeq reagent kit v2 with 2 × 250 cycles (Illumina, Inc.). A final base call validation of the PacBio contigs was performed using MiSeq reads trimmed using a quality score threshold of 20 or higher (≥Q20) and the reference assembler within Geneious software v11.1 (Biomatters, Ltd., Auckland, New Zealand). The final coverage for each

Received 5 June 2018 **Accepted** 6 June 2018 **Published** 12 July 2018

Citation Parker CT, Cooper KK, Huynh S, Smith TP, Bono JL, Cooley M. 2018. Genome sequences of eight Shiga toxin-producing *Escherichia coli* strains isolated from a produce-growing region in California. Microbiol Resour Announc 7:e00807-18. https://doi.org/10.1128/MRA.00807-18.

Editor Frank J. Stewart, Georgia Institute of Technology

Address correspondence to Craig T. Parker, craig.parker@ars.usda.gov.

TABLE 1 Accession numbers and genome metrics of 8 Shiga toxin-producing *Escherichia coli* strains from a produce-growing region in California

Strain name	Serotype	Chromosomal GenBank accession no.	No. of *stx* phages	LEE region	Chromosome size (bp)	Plasmid size (bp) (GenBank accession no.)
RM8426	O26:H11	CP028116	1	Yes	5,648,177	90,123 (CP028115)
RM10386	O26:H11	CP028126	2	Yes	5,785,882	83,012 (CP028124) 98,899 (CP028125)
RM10042	O43:H2	CP028122	2	No	5,057,506	8,120 (CP028123) 86,617 (CP028120) 109,466 (CP028121)
RM8385	O103:H11	CP028112	1	Yes	5,631,913	6,673 (CP028114) 94,220 (CP028113)
RM9322	O111:H8	CP028117	1	Yes	5,198,840	7,072 (CP028119) 78,469 (CP028118)
RM10466	O113:H21	CP028381	3	No	5,044,460	67,750 (CP028383) 160,675 (CP028382)
RM8352	O121:H19	CP028110	1	Yes	5,391,064	83,211 (CP028111)
RM9872	O145:H28	CP028379	1	Yes	5,385,924	89,518 (CP028380)

strain was >200×. Protein-, rRNA-, and tRNA-coding genes were annotated using the NCBI Prokaryotic Genomes Automatic Annotation Pipeline (PGAAP) (https://www.ncbi.nlm.nih.gov/genomes/static/Pipeline.html).

The accession numbers, serotypes, and assembly metrics for each complete genome are listed in Table 1. Seven serotypes are represented among these agriculturally relevant STEC genomes, including the first complete STEC O43 genome (strain RM10042) based on BLAST analysis. Features from these STEC genomes include one to three *stx*-containing prophages per chromosome. It should be noted that not all genomes possess the locus for enterocyte effacement (LEE) pathogenicity island. The sequencing data of these STEC strains will add to our understanding of STEC genome organization, phylogeny, and references for outbreak traceback analyses.

Data availability. The whole-genome sequences have been deposited in GenBank under the accession numbers listed in Table 1.

ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Agriculture, Agricultural Research Service CRIS projects 2030-42000-051-00D and 2030-42000-050-00D.

We thank Robert Mandrell, Anne Bates, and Mary Chapman for the curation of the Produce Safety and Microbiology Research Unit bacterial strain collection.

REFERENCES

1. Majowicz SE, Scallan E, Jones-Bitton A, Sargeant JM, Stapleton J, Angulo FJ, Yeung DH, Kirk MD. 2014. Global incidence of human Shiga toxin-producing *Escherichia coli* infections and deaths: a systematic review and knowledge synthesis. Foodborne Pathog Dis 11:447–455. https://doi.org/10.1089/fpd.2013.1704.
2. Batz MB, Hoffmann S, Morris JG Jr. 2012. Ranking the disease burden of 14 pathogens in food sources in the United States using attribution data from outbreak investigations and expert elicitation. J Food Prot 75:1278–1291. https://doi.org/10.4315/0362-028X.JFP-11-418.
3. Bettelheim KA. 2007. The non-O157 Shiga-toxigenic (verocytotoxigenic) *Escherichia coli*; under-rated pathogens. Crit Rev Microbiol 33:67–87. https://doi.org/10.1080/10408410601172172.
4. Brooks JT, Sowers EG, Wells JG, Greene KD, Griffin PM, Hoekstra RM, Strockbine NA. 2005. Non-O157 Shiga toxin-producing *Escherichia coli* infections in the United States, 1983–2002. J Infect Dis 192:1422–1429. https://doi.org/10.1086/466536.
5. Pohlenz TD, Zavadilova K, Ghosh A, Zurek L. 2018. Prevalence of Shiga-toxigenic *Escherichia coli* in house flies (Diptera: Muscidae) in an urban environment. J Med Entomol 55:436–439. https://doi.org/10.1093/jme/tjx225.
6. Alam MJ, Zurek L. 2004. Association of *Escherichia coli* O157:H7 with houseflies on a cattle farm. Appl Environ Microbiol 70:7578–7580. https://doi.org/10.1128/AEM.70.12.7578-7580.2004.
7. Berry ED, Wells JE, Bono JL, Woodbury BL, Kalchayanand N, Norman KN, Suslow TV, Lopez-Velasco G, Millner PD. 2015. Effect of proximity to a cattle feedlot on *Escherichia coli* O157:H7 contamination of leafy greens and evaluation of the potential for airborne transmission. Appl Environ Microbiol 81:1101–1110. https://doi.org/10.1128/AEM.02998-14.
8. Benjamin L, Atwill ER, Jay-Russell M, Cooley M, Carychao D, Gorski L, Mandrell RE. 2013. Occurrence of generic *Escherichia coli*, *E. coli* O157 and *Salmonella* spp. in water and sediment from leafy green produce farms and streams on the Central California coast. Int J Food Microbiol 165:65–76. https://doi.org/10.1016/j.ijfoodmicro.2013.04.003.
9. Jay MT, Cooley M, Carychao D, Wiscomb GW, Sweitzer RA, Crawford-Miksza L, Farrar JA, Lau DK, O'Connell J, Millington A, Asmundson RV, Atwill ER, Mandrell RE. 2007. *Escherichia coli* O157:H7 in feral swine near spinach fields and cattle, Central California coast. Emerg Infect Dis 13:1908–1911. https://doi.org/10.3201/eid1312.070763.
10. FDA. 2010. Federal and state officials confirm link between bagged Romaine lettuce and *E. coli* O145 illness outbreak. https://www.foodsafety.gov/news/educators/ednetmay2010.html#_Federal_and_State.

11. Askar M, Faber MS, Frank C, Bernard H, Gilsdorf A, Fruth A, Prager R, Höhle M, Suess T, Wadl M, Krause G, Stark K, Werber D; HUS Investigation Team. 2011. Large and ongoing outbreak of haemolytic uraemic syndrome, Germany, May 2011. Euro Surveill 16:19883. https://www.eurosurveillance.org/content/10.2807/ese.16.21.19878-en.
12. Centers for Disease Control and Prevention (CDC). 2018. Multistate outbreak of *E. coli* O157:H7 infections linked to romaine lettuce. Centers for Disease Control and Prevention, Atlanta, GA. https://www.cdc.gov/ecoli/2018/o157h7-04-18/index.html.
13. Centers for Disease Control and Prevention (CDC). 2006. Ongoing multistate outbreak of *Escherichia coli* serotype O157:H7 infections associated with consumption of fresh spinach–United States, September 2006. MMWR Morb Mortal Wkly Rep 55:1045–1046.
14. Centers for Disease Control and Prevention (CDC). 2017. Multistate outbreak of Shiga toxin-producing *Escherichia coli* O157:H7 infections linked to leafy greens (final update). Centers for Disease Control and Prevention, Atlanta, GA. https://www.cdc.gov/ecoli/2017/o157h7-12-17/index.html.
15. Cooley MB, Jay-Russell M, Atwill ER, Carychao D, Nguyen K, Quiñones B, Patel R, Walker S, Swimley M, Pierre-Jerome E, Gordus AG, Mandrell RE. 2013. Development of a robust method for isolation of Shiga toxin-positive *Escherichia coli* (STEC) from fecal, plant, soil and water samples from a leafy greens production region in California. PLoS One 8:e65716. https://doi.org/10.1371/journal.pone.0065716.
16. Miller WG, On SLW, Wang G, Fontanoz S, Lastovica AJ, Mandrell RE. 2005. Extended multilocus sequence typing system for *Campylobacter coli*, *C. lari*, *C. upsaliensis*, and *C. helveticus*. J Clin Microbiol 43:2315–2329. https://doi.org/10.1128/JCM.43.5.2315-2329.2005.
17. Miller WG, Yee E, Chapman MH, Smith TPL, Bono JL, Huynh S, Parker CT, Vandamme P, Luong K, Korlach J. 2014. Comparative genomics of the *Campylobacter lari* group. Genome Biol Evol 6:3252–3266. https://doi.org/10.1093/gbe/evu249.
18. PacBio. 2015. Procedure and checklist: 20 kb template preparation using BluePippin size-selection system. https://www.pacb.com/wp-content/uploads/2015/09/Procedure-Checklist-20-kb-Template-Preparation-Using-BluePippin-Size-Selection.pdf.